

# Statistical Methods for the $\beta$ -Binomial Model in Teratology

by Eiji Yamamoto<sup>1</sup> and Takemi Yanagimoto<sup>2</sup>

The  $\beta$ -binomial model is widely used for analyzing teratological data involving littermates. Recent developments in statistical analyses of teratological data are briefly reviewed with emphasis on the model. For statistical inference of the parameters in the  $\beta$ -binomial distribution, separation of the likelihood introduces an innovation in likelihood inference. This leads to reducing biases of estimators and also to improving accuracy of empirical significance levels of tests. Separate inference of the parameters can be conducted in a unified way.

## Introduction

Because teratological data include observations on fetuses from the same litter, binary responses have litter effects that cause overdispersion against the binomial model. By taking account the litter structure, several statistical models have been introduced, and many of their inference procedures have been proposed and improved. Reviews of this subject were presented in Haseman and Kupper (1) and in Krewski et al. (2). In the next section, we give a brief review of recent developments for statistical inference of the semiparametric model and the parametric model in the teratological data analysis and especially that of the  $\beta$ -binomial model. Then we review our recent work on modifications for the moment estimators of the parameters in the model.

Recent developments for likelihood inference emphasize advantages of separation of the likelihood (3). We will apply separate likelihood inference for the  $\beta$ -binomial population in the third section of this paper for expectation of improvement of the usual likelihood inference. A simulation study is conducted for examining performance of the applied inference procedures. In the final section, we discuss unsolved problems and future studies on the  $\beta$ -binomial model.

## Review of Teratological Data Analysis

### General View

For the test of the difference between prevalence rates in two samples, Gladen (4) proposed the jackknife method. On the assumption of the first two moments modeled on the litter structure, Williams (5) proposed the quasi-likelihood method for the dose-response regression analysis. On the other hand, the

binomial sampling error model was generalized for litter effects as the following parametric models. Williams (6) introduced the  $\beta$ -binomial model in the teratological data analysis. He assumed a  $\beta$  distribution between prevalence rates of litters. Kupper and Haseman (7) introduced the correlated binomial model by considering the correlation between two binary responses within the same litter. A different approach was used by Ochi and Prentice (8). In their model, binary responses within the same litter are defined according to whether the corresponding components of a multivariate normal variate with common mean, variance, and correlation exceed a common threshold. The usual likelihood methods for inference of the parameters have been used in the above models.

Among the existing models, the most important one is the  $\beta$ -binomial model. This model has been used widely in the analysis of teratological data and has been studied by many biostatisticians (9,10). Recent topics for the  $\beta$ -binomial model are concerned with the regression analysis and the incorporation of historical control data.

Kupper et al. (11) considered the fitting of a logistic dose-response curve to litter proportions in a  $\beta$ -binomial sampling error model. They showed from their simulation study that the maximum likelihood estimates (MLEs) of regression coefficients are seriously biased if the intralitter correlation is falsely assumed to be homogeneous across all dose groups. A simple explanation of the source of these large biases was given by Williams (12), and the theoretical aspect was discussed by Yamamoto and Yanagimoto (13).

Incorporating historical controls to a current toxicological experiment is another attractive topic. Tarone (14) assumed that the prevalence rate of the current control varies according to a  $\beta$  distribution. Hoel and Yanagawa (15) constructed a conditional test given the fixed number of responses in the current control group. In applications to actual data, estimates of the parameters in the  $\beta$  distribution are necessary, which might be obtained from the historical control data distributed in a  $\beta$ -binomial distribution. Recently, Prentice et al. (16) conducted a non-Bayes approach to incorporating the historical control data. They assumed that the historical controls follow a  $\beta$ -binomial distribution and the current experiment data follow a binomial or a

<sup>1</sup>Department of Applied Mathematics, Okayama University of Science, 1-1 Ridai-cho, Okayama 700, Japan.

<sup>2</sup>Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan.

Address reprint requests to E. Yamamoto, Department of Applied Mathematics, Okayama University of Science, 1-1 Radai-cho, Okayama 700, Japan.

This paper was presented at the International Biostatistics Conference on the Study of Toxicology that was held May 13-25, 1991, in Tokyo, Japan.

$\beta$ -binomial distribution. Inference of the parameters in an applied model is based on the joint likelihood of the historical and the current data.

## $\beta$ -Binomial Distribution

Let  $n_i$  denote the size of the  $i$ th litter ( $i = 1, \dots, m$ ), and let  $x_i$  denote the number of affected fetuses. The number  $x_i$  is assumed to be distributed in a binomial distribution  $Bi(n_i, p_i)$  for a fixed prevalence rate,  $p_i$ . In addition, the prevalence rate,  $p_i$ , is assumed to follow a  $\beta$  distribution, say  $BB(n_i, \pi, \phi)$ , which has the following probability function:

$$p(x; n, \pi, \phi) = \binom{n}{x} \frac{\prod_{r=0}^{x-1} (\pi + r\theta) \prod_{r=0}^{n-x-1} (1 - \pi + r\theta)}{\prod_{r=0}^{n-1} (1 + r\theta)},$$

where  $x = 0, 1, \dots, n$ ;  $\theta = \phi/(1 - \phi)$ ,  $0 \leq \phi < 1$ ,  $0 < \pi < 1$ . The mean,  $\mu$ , and the variance,  $\sigma^2$ , of the distribution are  $n\pi$  and  $n\pi(1 - \pi)[1 + (n - 1)\phi]$ , respectively. The parameter  $\pi$  represents the incidence rate of binary responses, and the parameter  $\phi$  represents the positive correlation coefficient between two binary responses. The marginal point  $\phi = 0$  means the binomial distribution so that  $\phi$  is regarded as an index of overdispersion against the binomial model. Prentice (17) noted that the  $\beta$ -binomial distribution formally covers underdispersion to a limited extent.

Here we note that the  $\beta$ -binomial distribution does not have favorable analytical properties. The distribution has the antimode when  $\phi$  is large. Explicitly, when  $(n + 1)\pi - (n - 1)\theta - 1 < 0$  and  $(n + 1)\pi - (n - 1)(1 - \theta) - 1 > 0$ , the point  $[(n + 1)(\pi - \theta)/(1 - 2\theta)]$  is the antimode, where  $[\ ]$  is the Gauss symbol. Classification in various shapes of the probability function by fixed-parameter values are illustrated in Figure 1. The distribution is not a member of the exponential family nor the exponential dispersion model (18). The parameters  $\pi$  and  $\phi$  are not orthogonal (19), except when  $\phi = 0$ . It is not reproductive; that is, the sample sum has a complex probability function. The MLE of the mean,  $\mu$ , is not the sample mean and cannot be expressed in a closed form.

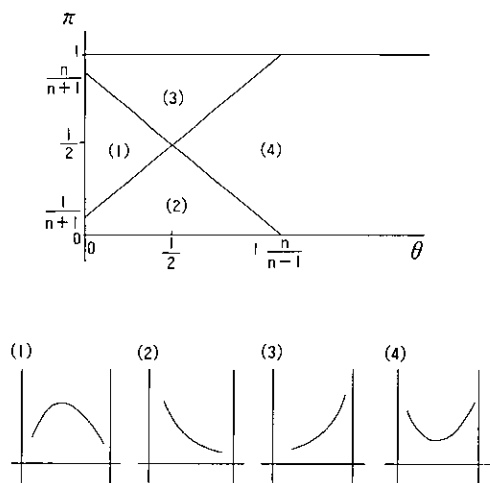


FIGURE 1. Division of the parameter space corresponding to classification in the four shapes of the probability function in the  $\beta$ -binomial distribution.

## Statistical Inference of $(\pi, \phi)$

Let us consider statistical inference of the parameters  $(\pi, \phi)$ . For simplicity we assume that the number of fetuses,  $n_i$ , is common among litters. Because of the above unfavorable properties, the two traditional estimation methods, the method of moments and the maximum likelihood method, have been used routinely. A few attempts to improve estimators have been made. Kleinman (20) claimed the superiority of the moment estimator of  $\pi$  with proper weights. Tamura and Young (21) proposed the use of a stabilizer for the usual moment estimator of  $\phi$ . Crowder (22) suggested good performance of the conditional MLE of  $\phi$  fixed the sample mean,  $\bar{x}$ , where he approximated the distribution of  $\bar{x}$  by a normal distribution.

For the test of the null hypothesis  $\phi = 0$ , the traditional asymptotic likelihood ratio test (LRT) theory has been falsely applied in spite of the fact that  $\phi = 0$  is the marginal point of the parameter range. Paul et al. (23) claimed that the LRT statistic is asymptotically distributed in the 50:50 mixture of the degenerate distribution at zero and the chi-square distribution with 1 degree of freedom (df) under the null hypothesis. The  $C(\alpha)$ -test proposed by Tarone (24) and recommended by Paul et al. (23) is able to test this marginal point without any difficulty. Accuracy of the empirical significance level can be improved by using an alternative asymptotic distribution of the test statistic under the assumption of large litter sizes proposed by Kim and Margolin (25). As Prentice (17) noted, in the extended  $\beta$ -binomial distribution, the point zero can be treated as an inner point of the parameter range so that the traditional LRT theory can be applied correctly.

## Method of Moments

The moment estimator of  $\pi$  is  $\hat{\pi} = \bar{x}/n$ , which is unbiased and has a potential efficiency (20). The moment estimator of  $\phi$  is denoted by

$$\hat{\phi}_{mo} = \frac{ns^2 - \bar{x}(n - \bar{x})}{(n - 1)\bar{x}(n - \bar{x})}$$

where  $s^2 = \Sigma(x_i - \bar{x})^2/(m - 1)$ , which is known to have an asymptotic positive bias. The estimator  $\hat{\phi}_{mo}$  is given as the root of the estimating equation (26)

$$g_b(\underline{x}; \phi) = ns^2 - \bar{x}(n - \bar{x}) - \phi(n - 1)\bar{x}(n - \bar{x}) = 0,$$

which is not unbiased, that is,  $E[g_b(\underline{x}; \phi)] \neq 0$ . This comes from the fact that  $\bar{x}(n - \bar{x})$  is not an unbiased estimator of  $n^2\pi(1 - \pi)$ . Yanagimoto and Yamamoto (27) claimed that removing the bias from an estimating equation for a usual moment estimator leads to better performances in many examples appearing in the actual statistical analyses. An unbiased estimating equation for  $\phi$

$$g(\underline{x}; \phi) = (mn - 1)s^2 -$$

$$m\bar{x}(n - \bar{x}) - \phi(n - 1)(m\bar{x}(n - \bar{x}) + s^2) = 0,$$

gives a moment estimator

$$\hat{\phi} = \frac{(mn - 1)s^2 - m\bar{x}(n - \bar{x})}{(n - 1)(m\bar{x}(n - \bar{x}) + s^2)}. \quad (1)$$

This treatment reduces the bias and the mean square error. Yamamoto and Yanagimoto (28) made an extensive comparison of performance of  $\hat{\phi}$  and  $\hat{\phi}_{mo}$ .

Until now the method of moments focuses only on the estimation procedure and consequently does not attract our attention to the test procedure. By using an unbiased moment estimating equation  $g(\underline{x}; \theta) = 0$ , we can produce a following test statistic for the null hypothesis  $\theta = \theta_0$  defined by

$$T_{mo} = \frac{\int_{\theta_0}^{\bar{\theta}} g(\underline{x}; \theta) d\theta}{E_{st.} \text{ of } E(\int_{\theta_0}^{\bar{\theta}} g(\underline{x}; \theta) d\theta)}$$

where  $\bar{\theta}$  is the unbiased moment estimator given by  $g(\underline{x}; \theta) = 0$  and the denominator is an unbiased moment estimator for the expectation of  $\int_{\theta_0}^{\bar{\theta}} g(\underline{x}; \theta) d\theta$ . Note that the test statistic  $T_{mo}$  can be explained as a signal-noise ratio under the hypothesis. Applying this test procedure to the  $\beta$ -binomial case, the test statistic for the null hypothesis  $\pi = \pi_0$  is given by

$$T_{mo} = \frac{(\bar{x} - n\pi_0)^2}{s^2/n}$$

which is the square of the well known  $t$ -test statistic. The test for  $\phi$  needs complicated calculations of the third and the fourth moments, so we do not pursue the moment test procedure for  $\phi$  any further.

## Innovation in Likelihood Inference

### Outline of Principle

Recent developments for likelihood inference of the mean,  $\mu$ , and the dispersion parameter,  $\theta$ , put emphasis on the advantage of separation of the likelihood (3), which is based on factorization of the density function of a sample  $\underline{x}$ ;

$$f(\underline{x}; \mu, \theta) = f_m(t; \mu, \theta) f_c(\underline{x}, \theta | t), \quad (2)$$

where  $t$  is the sample sum or the sample mean. The marginal density is used as the likelihood for inference of  $\mu$ , and the conditional density is used as the likelihood for  $\theta$ . The maximum likelihood estimation procedure gives the marginal MLE  $\hat{\mu}$  and the conditional MLE  $\hat{\theta}$ . In this factorization (Equation 2),  $\hat{\mu}$  is also the usual MLE. Yanagimoto and Yamamoto (29) proposed a modified likelihood ratio test statistic for the null hypothesis  $\mu = \mu_0$  as

$$T_m = 2 \ln \frac{f_m(t; \hat{\mu}, \hat{\theta})}{f_m(t; \mu_0, \hat{\theta})}.$$

Note that the conditional MLE  $\hat{\theta}$  is used both in the numerator and the denominator, and the test statistic is compared exactly or approximately with the  $F$  distribution with appropriate degrees of freedom. We will call  $T_m$  the marginal LRT for  $\mu$  hereafter. In the normality case,  $T_m$  is just equal to the square of the  $t$ -test statistic. On the other hand, the usual LRT statistic in this case is expressed as  $n \log [1 + T_m/(n - 1)]$ . Favorable per-

formances and many successful examples of the marginal LRT were presented in Yanagimoto and Yamamoto (29).

### Inference Procedures

Though the  $\beta$ -binomial distribution cannot be factored into the form of Equation 2, we can expect that the application of separation of the likelihood leads to improving usual likelihood inference. Inference of  $\pi$  is based on the marginal density of the sample sum ( $t = \sum x_i$ ) and that of  $\phi$  is based on the conditional density fixed  $t$ . Following the principle outlined above, the joint density of  $\underline{x}$  is separated in

$$f(\underline{x}; \pi, \phi) = f_m(t; \pi, \phi) f_c(\underline{x}; \pi, \phi | t).$$

Unfortunately, the marginal density  $f_m$  is of a complicated form, and also  $\pi$  remains in the conditional density  $f_c$ . Therefore, we will evaluate the former by the following approximation:

The first and second moments of  $t$  are given by

$$\begin{aligned} \mu_t &= E(t) = N\pi \\ \sigma_t^2 &= V(t) = N\pi(1 - \pi)(1 + (n - 1)\phi) \\ &= N\pi(1 - \pi)(1 + (N - 1)\phi') \end{aligned}$$

where  $N = mn$  and  $(N - 1)\phi' = (n - 1)\phi$ . Crowder (22) applied a normal distribution for a candidate of the approximated distribution of  $t$ . We propose here the use of a  $\beta$ -binomial distribution as a more reasonable candidate, because the distribution of  $t$  is skewed. Especially the  $\beta$ -binomial approximation has merits such that the sample distribution is discrete and closed in the  $\beta$ -binomial family. Fixing the first two moments of  $t$ , the above two approximated distributions have the following probability density and function:

$$f_m(t; \pi, \phi) = \frac{1}{\sqrt{2\pi}\sigma_t} \exp - \frac{(t - \mu_t)^2}{2\sigma_t^2}, \quad (3)$$

$$f_m(t; \pi, \phi) = \binom{N}{t} \frac{\prod_{r=0}^{t-1} (\pi + r\theta') \prod_{r=0}^{N-t-1} (1 - \pi + r\theta')}{\prod_{r=0}^N (1 - \pi + r\theta')}, \quad (4)$$

respectively, with  $\theta' = \phi'/(1 - \phi')$ . Consequently, the conditional density is approximated by

$$f_c(\underline{x}; \pi, \phi | t) = \frac{f(\underline{x}; \pi, \phi)}{f_m(t; \pi, \phi)},$$

where  $f_m(t; \pi, \phi)$  is the density (Equation 3 or 4).

By using the above approximations to the marginal likelihood for  $\pi$  and the conditional likelihood for  $\phi$ , we will introduce the following combination of estimators and test procedures in a unified way. a) Estimation of  $\pi$ : we use the moment estimator  $\hat{\pi} = \bar{x}/n$ . b) Estimation of  $\phi$ : the conditional MLE  $\hat{\phi}$  is given by the approximated conditional likelihood fixed  $t$  defined by

$$l_c(\phi) = f_c(\underline{x}; \hat{\pi}, \phi | t). \quad (5)$$

c) Test of  $\pi$ : the marginal LRT for the hypothesis  $\pi = \pi_0 (> 0)$  is proposed as

$$T_m = 2 \ln \frac{f_m(t; \hat{\pi}, \hat{\phi})}{f_m(t; \pi_0, \hat{\phi})},$$

which is compared with the  $F$  distribution with df  $(1, m-1)$ .

d) Test of  $\phi$ : the conditional LRT for the hypothesis  $\phi = \phi_0 (> 0)$  is proposed as

$$T_c = 2 \ln \frac{f_c(\underline{x}; \hat{\pi}, \hat{\phi}|t)}{f_c(\underline{x}; \hat{\pi}, \phi_0|t)},$$

which is compared with the chi-square distribution with df 1.

When we approximate the marginal distribution of  $t$  by the  $\beta$ -binomial distribution, the conditional likelihood of  $\phi$  in Equation 5 has the following derivative at the point  $\phi = 0$ .

$$\frac{d}{d\phi} l_c(\phi)|_{\phi=0} = \frac{Nn(m-1)}{2(N-1)} \left[ \frac{m(N-1)}{t(N-t)} s^2 - 1 \right].$$

Therefore, the condition  $\hat{\phi} > 0$  is equivalent with that  $\tilde{\phi} > 0$  in Equation 1.

## Simulation Study

To examine performance of the proposed estimators of  $\phi$  and the test procedures and  $\pi$  and  $\phi$ , we conducted a simulation study. The selected values of the incidence rate  $\pi$  are 0.05, 0.2, and 0.4 as small, moderate, and large values, respectively. The large value is set for a study of the behaviors of estimates out of the dependence on the constraint such that the estimates of  $\phi$  should be non-negative. The dispersion parameter  $\phi$  is selected suitably at each  $\pi$  level in 0.05 through 0.4. The  $\beta$ -binomial random numbers were generated by the IMSL package, and the MLE was obtained by the program in Smith (30). The size of effective iteration is 10,000.

**Estimation of  $\phi$ .** The estimators of  $\phi$  in the study are the MLE, the unbiased moment estimator  $\hat{\phi}$ , and the two conditional MLEs based on the normal approximation to the marginal distribution of  $t$ , say  $\hat{\phi}_N$ , and the  $\beta$ -binomial approximation, say  $\hat{\phi}_B$ . In our simulations, the cases where estimates of  $\phi$  were indeterminate or took the value 1 occurred rarely, and they were not counted in effective iteration. When an estimate took a negative value, it was regarded to take zero. For each estimator, we calculated the median bias, which is defined by the median deviation from the true value, and the MSE. The median bias is used because it is hardly influenced by the non-negative constraint to estimates.

The results are shown in Tables 1 and 2. Noteworthy findings in Table 1 are that the median biases of the two conditional MLEs are about half those of the MLEs. The unbiased moment estimator decreases by 1 when  $\pi$  is large. Table 2 indicates that the MSEs of the four estimators are comparable to each other. Summarizing results of the simulations, we conclude that the conditional MLEs perform the best, and the unbiased moment estimator is superior to the MLE.

**Table 1. Median biases of estimators for the parameter  $\phi$  (litter size = 10, number of animals = 20, number of iterations = 10,000).**

$\pi$	$\phi$	MLE	$\hat{\phi}_N$	$\hat{\phi}_B$	$\hat{\phi}$
0.05	0.01	-0.0100	-0.0077	-0.0077	-0.0083
	0.02	-0.0138	-0.0056	-0.0055	-0.0100
	0.04	-0.0147	-0.0057	-0.0056	-0.0140
	0.08	-0.0191	-0.0067	-0.0058	-0.0259
	0.16	-0.0340	-0.0095	-0.0109	-0.0438
0.2	0.01	-0.0091	-0.0038	-0.0037	-0.0038
	0.02	-0.0104	-0.0044	-0.0044	-0.0052
	0.04	-0.0115	-0.0049	-0.0049	-0.0059
	0.08	-0.0136	-0.0059	-0.0058	-0.0087
	0.16	-0.0170	-0.0070	-0.0068	-0.0109
0.4	0.2	-0.0172	-0.0086	-0.0085	-0.0047
	0.25	-0.0161	-0.0075	-0.0073	-0.0026
	0.3	-0.0153	-0.0069	-0.0066	-0.0025
	0.35	-0.0166	-0.0086	-0.0083	-0.0024
	0.4	-0.0166	-0.0089	-0.0086	-0.0011

MLE, maximum likelihood estimate.

**Table 2. Mean square errors of estimators for the parameter  $\phi$  (litter size = 10, number of animals = 20, number of iterations = 10,000).**

$\pi$	$\phi$	MLE	$\hat{\phi}_N$	$\hat{\phi}_B$	$\hat{\phi}$
0.05	0.01	0.0010	0.0016	0.0016	0.0009
	0.02	0.0015	0.0021	0.0022	0.0013
	0.04	0.0026	0.0035	0.0036	0.0023
	0.08	0.0053	0.0068	0.0071	0.0050
	0.16	0.0125	0.0150	0.0173	0.0129
0.2	0.01	0.0007	0.0009	0.0009	0.0008
	0.02	0.0009	0.0011	0.0011	0.0011
	0.04	0.0015	0.0017	0.0017	0.0017
	0.08	0.0028	0.0029	0.0030	0.0030
	0.16	0.0057	0.0057	0.0059	0.0063
0.4	0.2	0.0055	0.0054	0.0054	0.0058
	0.25	0.0067	0.0066	0.0066	0.0070
	0.3	0.0076	0.0074	0.0074	0.0080
	0.35	0.0085	0.0082	0.0082	0.0088
	0.4	0.0091	0.0087	0.0090	0.0095

MLE, maximum likelihood estimate.

**Test of  $\pi$ .** From the above results, it looks better to use the two conditional MLEs of  $\phi$  for the marginal LRT for  $\pi$ . Then the test procedures for  $\pi$  in the present study consist of the usual LRT, the squared  $t$ -test, which is derived by the method of moments, and the two marginal LRTs based on the normal, say  $T_{mN}$ , and the  $\beta$ -binomial approximation of  $t$ , say  $T_{mB}$ . The parameter values of  $(\pi, \phi)$  are same as in the situation for the estimation of  $\pi$ . For each test procedure, empirical significant levels for the nominal 5 and 1% levels are examined.

The results are given in Tables 3 and 4, showing that the usual LRT overstates statistical significance when  $\phi$  is larger than 0.04. The squared  $t$ -test overstates all the parameter values and has stable empirical levels when  $\pi$  is large. The approximations by the normal distribution and the  $\beta$ -binomial distribution to the marginal distribution of  $t$  have resulted in the similar performance for the estimation of  $\phi$ , but  $T_{mN}$  shows differences from  $T_{mB}$ . The empirical levels of  $T_{mN}$  are unstable, and their range is wider than that of the usual LRT.

On the other hand,  $T_{mB}$  has smaller empirical levels than

the usual LRT for all the parameter values. The understatement of  $T_{mB}$  can be improved by using the chi-square distribution in place of the  $F$  distribution to yield the critical values when an estimate  $\hat{\phi}$  is 0. In the right-hand columns of Tables 3 and 4, modified empirical levels are given. The improvement in the case of small  $\phi$  looks satisfactory. Such a modification is supported by the fact that when  $\phi = 0$ , the  $\beta$ -binomial distribution becomes the binomial distribution so that the chi-square test is more appropriate when  $\phi = 0$ . Recall that a negative estimate  $\hat{\phi}$  is changed to zero.

**Test of  $\phi$ .** The test for the hypothesis  $\phi = \phi_0 (> 0)$  is performed by the two conditional LRTs based on the normal, say  $T_{cN}$ , and the  $\beta$ -binomial, say  $T_{cB}$ , approximation to the marginal distribution of  $t$  in comparison to the usual LRT. The fixed parameters of  $(\pi, \phi)$  are also same as the simulation for the estimation of  $\pi$ . For the three LRT procedures, empirical significance levels for the nominal 5 and 1% levels are examined, which are summarized in Tables 5 and 6. These results show that in the small  $\phi_0$  cases the empirical levels of the conditional LRTs are about half of the nominal levels. It is our understanding that this phenomenon is due to the constraint that  $\hat{\phi}$  is non-negative. Therefore, we should conduct the one-sided test against the alternative  $\phi > \phi_0$  is small, which means that we might compare each of the empirical levels to half of the corresponding nominal level.

Note that the one-sided  $U$  test can be produced by the signed LRT (31) as

$$U_c = \text{sign}(\hat{\phi} - \phi_0) \sqrt{T_c},$$

and it is applicable even when  $\phi_0 = 0$ . Empirical significant levels given by the one-sided  $U$  test of above the three test are given in Tables 5 and 6. the signed LRT against the one-sided

**Table 3. Empirical significance levels of tests for the parameter  $\pi$  with 5% nominal level**  
(litter size = 10, number of animals = 20, number of iterations = 10,000).

$\pi$	$\phi$	LRT	$t^2$ -Test	$T_{mN}$	$T_{mB}$	$T_{mB}$
		$\chi^2$	$F$	$F$	$F$	$\chi^2 + F$
0.05	0.01	4.23	5.99	6.72	3.44	4.03
	0.02	4.74	6.29	7.39	3.88	4.39
	0.04	5.61	7.92	9.54	4.71	5.01
	0.08	6.54	9.03	11.03	5.59	5.67
	0.16	8.37	12.07	13.63	7.20	7.21
0.2	0.01	4.43	5.45	3.72	3.20	4.13
	0.02	4.81	5.35	4.08	3.68	4.36
	0.04	5.14	5.48	4.74	4.20	4.65
	0.08	5.71	5.81	5.72	4.86	5.12
	0.16	5.82	6.65	6.84	5.22	5.26
0.4	0.01	4.72	5.20	3.94	3.37	4.19
	0.02	5.01	5.32	4.36	3.74	4.46
	0.04	5.68	5.37	5.03	4.54	5.02
	0.16	6.10	5.50	5.81	5.17	5.20
	0.2	5.88	5.13	5.42	4.81	4.82
	0.25	5.69	5.02	5.63	4.72	4.72
	0.3	5.92	5.34	5.83	4.88	4.88
	0.35	5.61	5.05	5.63	4.69	4.69
	0.4	5.63	5.35	5.96	4.83	4.83

LRT, likelihood ratio test.

**Table 4. Empirical significance levels of tests for the parameter  $\pi$  with 1% nominal level**  
(litter size = 10, number of animals = 20, number of iterations = 10,000).

$\pi$	$\phi$	LRT	$t^2$ -Test	$T_{mN}$	$T_{mB}$	$T_{mB}$
		$\chi^2$	$F$	$F$	$F$	$\chi^2 + F$
0.05	0.01	1.24	2.47	2.42	0.50	1.12
	0.02	1.56	2.80	2.79	0.51	1.35
	0.04	2.19	3.78	3.76	0.77	1.85
	0.08	2.99	5.02	5.01	1.42	2.69
	0.16	5.29	8.09	8.08	2.80	4.83
0.2	0.01	0.80	1.02	0.64	0.40	0.74
	0.02	0.99	1.33	0.85	0.52	0.91
	0.04	1.08	1.35	1.10	0.62	0.90
	0.08	1.32	1.64	1.50	0.87	1.02
	0.16	1.43	1.98	2.02	1.22	1.25
0.4	0.01	0.81	1.16	0.55	0.44	0.74
	0.02	0.92	1.07	0.64	0.57	0.80
	0.04	1.19	1.16	0.88	0.70	0.93
	0.08	1.04	0.88	0.89	0.67	0.78
	0.16	1.31	1.19	1.31	1.01	1.04
	0.2	1.37	1.28	1.34	1.08	1.08
	0.25	1.18	1.18	1.26	0.96	0.96
	0.3	1.21	1.18	1.32	0.96	0.96
	0.35	1.17	1.19	1.35	0.97	0.97
	0.4	1.18	1.17	1.36	0.88	0.88

LRT, likelihood ratio test.

**Table 5. Empirical significance levels of tests for the parameter  $\phi$  with 5% nominal level**  
(litter size = 10, number of animals = 20, number of iterations = 10,000).

$\pi$	$\phi$	LRT	$T_{cN}$	$T_{cB}$
0.05	0.00	(3.43) <sup>a</sup>	(3.82)	(3.90)
	0.01	1.25 (2.92)	2.29 (4.23)	2.31 (4.23)
	0.02	1.27 (2.58)	2.27 (4.62)	2.27 (4.62)
	0.04	1.61 (2.84)	2.67 (4.75)	2.75 (4.75)
	0.08	3.65	4.68	4.68
	0.16	5.56	7.63	7.73
0.2	0.00	(3.16)	(4.19)	(4.20)
	0.01	1.51 (3.10)	2.13 (4.04)	2.13 (4.04)
	0.02	1.43 (3.15)	2.07 (4.23)	2.08 (4.25)
	0.04	2.14 (3.15)	2.47 (4.30)	2.48 (4.33)
	0.08	5.28	4.87	4.86
	0.16	5.86	6.14	6.17
0.4	0.00	(3.18)	(4.54)	(4.55)
	0.01	1.57 (3.20)	2.12 (4.23)	2.13 (4.24)
	0.02	1.30 (3.18)	1.90 (4.37)	1.90 (4.37)
	0.04	2.04 (3.10)	2.12 (4.09)	2.14 (4.11)
	0.08	5.77	5.07	5.05
	0.16	5.86	5.26	5.25
	0.2	6.08	5.43	5.43
	0.25	6.04	5.64	5.63
	0.3	6.07	5.65	5.64
	0.35	6.30	5.97	5.96
	0.4	6.40	6.14	6.12

LRT, likelihood ratio test.

<sup>a</sup>Values in parentheses were obtained using a one-sided test.

alternative leads to better accuracy of empirical levels for the two conditional LRTs than the usual LRT.

In conclusion, the simulation study has shown that separate likelihood inference has the ability to innovate in statistical inference of the parameters  $(\pi, \phi)$  in the  $\beta$ -binomial distribution.

**Table 6. Empirical significance levels of tests for the parameter  $\phi$  with 1% nominal level (litter size = 10, number of animals = 20, number of iterations = 10,000).**

$\pi$	$\phi$	LRT	$T_{CN}$	$T_{CB}$
0.05	0.00	(0.52) <sup>a</sup>	(1.04)	(1.04)
	0.01	0.22 (0.48)	0.43 (0.90)	0.43 (0.90)
	0.02	0.30 (0.60)	0.61 (0.94)	0.61 (0.94)
	0.04	0.31 (0.63)	0.67 (1.31)	0.67 (1.31)
	0.08	0.23	0.70	0.70
	0.16	0.54	1.07	1.07
0.2	0.00	(0.57)	(0.84)	(0.84)
	0.01	0.17 (0.52)	0.30 (0.96)	0.30 (0.97)
	0.02	0.18 (0.53)	0.40 (0.70)	0.40 (0.70)
	0.04	0.28 (0.63)	0.48 (0.86)	0.50 (0.86)
	0.08	0.46	0.54	0.54
	0.16	1.33	1.36	1.37
0.4	0.00	(0.58)	(0.86)	(0.86)
	0.01	0.24 (0.49)	0.38 (0.77)	0.38 (0.77)
	0.02	0.23 (0.52)	0.36 (0.78)	0.36 (0.78)
	0.04	0.28 (0.53)	0.37 (0.72)	0.37 (0.72)
	0.08	0.69	0.65	0.65
	0.16	1.55	1.34	1.34
	0.2	1.45	1.30	1.30
	0.25	1.53	1.39	1.40
	0.3	1.37	1.40	1.39
	0.35	1.37	1.31	1.29
	0.4	1.46	1.34	1.35

LRT, likelihood ratio test.

<sup>a</sup>Values in parentheses were obtained using a one-sided test.

## Further Problems

In this paper we do not consider the heterogeneous litter-size case, two-sample problems, and the regression analysis. Notice that the inference procedures proposed above are derived from separation of the likelihood. We expect that this principle can be applied to these statistical problems successfully. For example, in two-sample problems, the estimation of a common  $\phi$  may be conducted by maximizing the conditional likelihood

$$\frac{p_1(\bar{x}, \hat{\pi}_1, \phi)p_2(\bar{y}, \hat{\pi}_2, \phi)}{p_{1m}(t_1, \hat{\pi}_1, \phi)p_{2m}(t_2, \hat{\pi}_2, \phi)}$$

where  $t_1 = \sum x_i$ ,  $t_2 = \sum y_i$ ,  $\hat{\pi}_1 = \bar{x}/n_1$ ,  $\hat{\pi}_2 = \bar{y}/n_2$ , and the marginal distributions of  $t_1$  and  $t_2$  are adjusted to  $\beta$ -binomial distributions, respectively. The  $t$ -test for the difference between two incidence rates with a common dispersion may be constructed by the signed marginal LRT

$$sgn(\bar{x} - \bar{y}) \sqrt{2 \ln \frac{p_{1m}(t_1, \hat{\pi}_1, \hat{\phi})p_{2m}(t_2, \hat{\pi}_2, \hat{\phi})}{p_{1m}(t_1, \hat{\pi}, \hat{\phi})p_{2m}(t_2, \hat{\pi}, \hat{\phi})}}$$

where  $\hat{\pi} = (m_1\hat{\pi}_1 + m_2\hat{\pi}_2)/(m_1 + m_2)$  and  $\hat{\phi}$  is the above conditional MLE. For regression problems, the construction of

inference procedures for regression coefficients looks more difficult but is worth future study.

The authors thank Byung Soo Kim and Barry H. Margolin for their helpful comments.

## REFERENCES

1. Haseman, J. K., and Kupper, L. L. Analysis of dichotomous response data from certain toxicological experiments. *Biometrics* 35: 281-293 (1979).
2. Krewski, D., Colin, D. and Yanagimoto, T. Statistical methods for the developmental toxicity studies. In: *Current Issues in Teratology* (K. Khera, H. Grice, and D. Clegg, Eds.), Springer-Verlag, New York, 1989, pp. 69-80.
3. Barndorff-Nielsen, O. E. *Parametric Statistical Models and Likelihood*. Springer-Verlag, New York, 1988.
4. Gladen, B. The use of the jackknife to estimate proportions from toxicological data in the presence of litter effects. *J. Am. Stat. Assoc.* 74: 278-283 (1979).
5. Williams, D. A. Extra-binomial variation in logistic linear models. *Appl. Stat.* 31: 144-148 (1982).
6. Williams, D. A. The analysis of binary responses from toxicological experiments involving reproduction and teratogenicity. *Biometrics* 31: 946-952 (1975).
7. Kupper, L. L., and Haseman, J. K. The use of a correlated binomial model for the analysis of certain toxicological experiments. *Biometrics* 34: 69-76 (1978).
8. Ochi, Y., and Prentice, R. L. Likelihood inference in a correlated probit regression model. *Biometrika* 71: 531-543 (1984).
9. Segreti, A. C., and Munson, A. E. Estimation of the median lethal dose when responses within a litter are correlated. *Biometrics* 37: 153-156 (1981).
10. Paul, S. R. Analysis of proportions of affected fetuses in teratological experiments. *Biometrics* 38: 361-370 (1982).
11. Kupper, L. L., Portier, C., Hogan, M. D., and Yamamoto, E. The impact of litter effects on dose-response modeling in teratology. *Biometrics* 42: 85-98 (1986).
12. Williams, D. A. Estimation bias using the beta-binomial distribution in teratology. *Biometrics* 44: 305-309 (1988).
13. Yamamoto, E., and Yanagimoto, T. Litter effects to dose-response curve estimation. *J. Jn. Stat. Soc.* 18: 97-106 (1988).
14. Tarone, R. E. The use of historical control information in testing for a trend in proportion. *Biometrics* 38: 215-220 (1982).
15. Hoel, D. G., and Yanagawa, T. Incorporating historical controls in testing for a trend in proportions. *J. Am. Stat. Assoc.* 81: 1095-1099 (1986).
16. Prentice, R. L., Smythe, R. T., Krewski, D., and Mason, M. On the use of historical control data to estimate dose response trends in quantal bioassay. *Biometrics* 48: 459-478 (1992).
17. Prentice, R. L. Binary regression using an extended beta-binomial distribution, with discussion of correction induced by covariate measurement errors. *J. Am. Stat. Assoc.* 81: 321-327 (1986).
18. Jorgensen, B. Exponential dispersion models. *J. R. Stat. Soc. B* 49: 127-162 (1987).
19. Cox, D. R., and Reid, N. Parameter orthogonality and approximated conditional inference, with discussion. *J. R. Stat. Soc. B* 49: 1-39 (1987).
20. Kleinman, J. C. Proportions with extraneous variance: single and independent samples. *J. Am. Stat. Assoc.* 68: 46-54 (1973).
21. Tamura, R. N., and Young, S. S. A stabilized moment estimator for the beta-binomial distribution. *Biometrics* 43: 813-824 (1987).
22. Crowder, M. J. Inference about the intraclass correlation coefficient in the beta-binomial ANOVA for proportions. *J. R. Stat. Soc. B* 41: 230-234 (1979).
23. Paul, S. R., Liang, K. Y., and Self, S. G. On testing departure from the binomial and multinomial assumptions. *Biometrics* 45: 231-236 (1989).
24. Tarone, R. E. Testing the goodness of fit of the binomial distribution. *Biometrika* 66: 585-590 (1979).
25. Kim, B. S., and Margolin, B. H. Testing goodness of fit of a multinomial model against overdispersed alternatives. *Biometrics* 48: 711-719 (1992).
26. Godambe, V. P. Conditional likelihood and unconditional optimum estimating equations. *Biometrika* 63: 277-284 (1976).
27. Yanagimoto, T., and Yamamoto, E. The role of unbiasedness in estimating equations. *Estimating Function* (V. P. Godambe, Ed.), Oxford University Press, Oxford, 1991, pp. 89-101.

28. Yamamoto, E., and Yanagimoto, T. Moment estimators for the beta-binomial distribution. *J. Appl. Stat.*, in press.
29. Yanagimoto, T. and Yamamoto, E. The likelihood ratio test of the mean in conditional inference. Technical Report of the Institute of Statistical Mathematics in Japan ISM-368 (1990).
30. Smith, D. M. Maximum likelihood estimation of the parameters of the beta-binomial distribution. *Appl. Stat.* 32: 196–204 (1983).
31. Barndorff-Nielsen, O. E. Inference on full or partial parameters based on the standardized signed log likelihood ratio. *Biometrika* 73: 307–322 (1986).